

Thomas Stäcker

Die Sammlung ist tot, es lebe die Sammlung! – Die digitale Sammlung als Paradigma moderner Bibliotheksarbeit

Zusammenfassung: Mit dem digitalen Wandel verändert sich der überlieferte Begriff der Sammlung. Während früher Bibliotheken ihre Bedeutung aus der Größe und Qualität ihrer Sammlung ableiteten, sind sie heute angesichts der zunehmend frei verfügbaren digitalen Quellen und Dokumente gezwungen umzudenken. Der Beitrag vertritt die Position, dass die Kultur des Besitzens einer Kultur des Teilens weichen muss. Nicht der exklusive Besitz eines Mediums entscheidet über die Qualität einer Bibliothek, sondern der nicht-exklusive Besitz. Dokumente und Daten werden aus der Logik des Mediums heraus in Zukunft durchgehend im Open Access angeboten werden. Das bedeutet aber nicht, dass sich das Sammeln erübrigt. Die Inadäquatheit, Disparität und Inhomogenität der analogen oder freien digitalen Angebote verlangt nach Transformation, Aggregation und Strukturierung bzw. Homogenisierung der Dokumente und Daten, um über die Herstellung der Maschinenlesbarkeit nach den FAIR Prinzipien attraktive Angebote von qualitativ hochwertigen digitalen Daten- und Dokumentsammlungen und Korpora für die mit digitalen Methoden arbeitende Forschung machen zu können. Für die Bibliothek eröffnen sich unter dem Begriff der <soCalled>Datafication</soCalled> wichtige neue Arbeitsfelder, die zwar die alte Sammlung obsolet machen, aber sie eine neue gewinnen lassen.

Schlüsselwörter: Digitale Sammlung, Forschungsdaten

The Collection is Dead Long Live the Collection – The Digital Collection as Paradigm of Modern Library Work

Abstract: The media change alters the traditional concept of the collection. While in earlier times libraries obtained their reputation by size and rareness of their collections, today in view of the increasing availability of free resources on the web they are forced to reevaluate this conviction. This paper argues that the culture of exclusiveness must give way to a culture of sharing. Not the exclusive, but the non-exclusive ownership will determine the quality of the library of the future. Due to the logic of the media documents and data will be provided Open Access. However, this

doesn't mean that collecting is superfluous. The inadequateness, disparity and inhomogeneity of analogue or open digital resources requires their transformation, aggregation and structuring or homogenizing in order to create qualitatively attractive offers of digital collections and corpora for digital research by making these resources machine-readable according to the FAIR principles. This opens up new fields of work for the library in that it provides new <soCalled>datafication</soCalled> services. The old collection will be obsolete, but the library will gain a new one instead.

Keywords: Digital collection, research data

Das Coverbild des jüngsten Buchs von Michael Knoche¹ zur Idee der Bibliothek und ihrer Zukunft zeigt den berühmten Rotundensaal der Herzogin Anna Amalia Bibliothek. Aber die Regale sind leer. Das liest sich wie ein Stigma, gegen das Knoche anschreibt. Die bibliothekarische wissenschaftliche Sammlung, so das Credo von Publikationen der letzten Jahre,² verliert angesichts der ubiquitär verfügbaren digitalen Werke an Bedeutung oder wandelt doch stark ihr Gesicht. Der sich abzeichnende Trend zu Open Access scheint diesen Eindruck noch zu verstärken. Spitzt man die Frage zu, dann erhebt sich die Frage: Was sammelt die Bibliothek noch, wenn alle Werke Open Access sind? Wie sieht dann dieser doch von allen Bibliotheken, die unter der Last unbezahlbar gewordenen Subskriptionsgebühren der Voldemorts der Verlagsbranche stöhnen, herbeigesehnte Zustand aus? Sammeln war gestern, doch was ist heute? Eine Antwort darauf versuchen Bibliothekare wie Lankes, der nicht ohne evangelikalen Unterton³ postuliert, dass es Bibliotheken weniger darum ginge „Bücher“ anzuhäufen als ihre Communities besser zu machen.⁴ Dazu sind Bücher allenfalls ein Hilfsmittel, aber kein Selbstzweck. So nötigt er uns, sich die Bibliothek allein mit Büchern vorzustellen. Es funktioniert nicht, so Lankes, denn das Wichtige seien nicht die Bücher, sondern die Bibliothekare.⁵

Diese ebenso verblüffende wie charmante Antwort täuscht aber darüber hinweg, dass sie ebenso für Lehrer, Professoren, Kindergärtner, Psychologen oder Pastoren gilt, die nach ihrem Selbstverständnis ihre „communities“ besser machen. Nun geht es um Wissensvermittlung. Da mag man vielleicht bei der einen oder anderen Berufssparte geteilter Meinung sein, doch gibt es in diesem Aspekt kaum eine *differentia specifica* des Bibliothekars. Es wäre denn auch eigenartig, wenn nun nach gut 2000 Jahren bibliothekarischer Sammlungsgeschichte, plötzlich das Paradigma der Sammlung als Irrweg

¹ Knoche (2017).

² Vgl. z.B. Ceynova (2015).

³ Lankes (2018) 341: „The irony of information literacy, indeed all of education, is that to make someone smarter, you must first get them to accept how ignorant they are.“ Das hat einen aufklärerischen Impuls, doch fehlt die Rückbindung an die Instanz, die allein über das Bessere entscheiden kann: die aufklärerische Vernunft.

⁴ Lankes (2016).

⁵ Lankes (2016) 16.

enttarnt und der Bibliothekar durch den digitalen Weckruf seine eigentliche Bestimmung gefunden hätte. Insofern sollte man sich, ehe man vorschnell die Sammlung als irrelevantes Beiwerk verabschiedet, solchen ansonsten sicher erfrischenden Thesen mit der Vorsicht eines an Jahren reifen und an vermeintlichen Paradigmenwechseln geläuterten Berufsstandes gegenüberreten, um nüchtern die sich ändernden Rahmenbedingungen zu analysieren und die Frage des Wertes und Bedeutung der Sammlung neu zu stellen.⁶

Den Begriff der Sammlung⁷ steht, wenn unter dem Gesichtspunkt des Digitalen betrachtet, in engem Zusammenhang mit dem Begriff der Forschungsdaten und wird auch regelmäßig in diesem Kontext diskutiert.⁸ Zugleich ist sichtbar, dass eine nähere Begriffsbestimmung, was eigentlich das Digitale ausmacht, unverzichtbar ist, wenn man dem Medium gerecht werden will. So fragt Wolfgang Ernst zurecht: „Sind Bibliotheken an die spezifische Materialität von Manuskript und Buch(druck) gebunden, oder gehen diese als Formate im digitalen Raum auf?“⁹ Digitale und analoge Medien besitzen zwar Gemeinsamkeiten, vor allem darin, dass sie Medien sind, doch zugleich sind sie auch grundsätzlich verschieden, und man sollte sich der jeweiligen Bedeutung und Funktionsweise versichern, indem man z.B. mit dem Kollektivautor Pédauque unter den drei Aspekte des digitalen Dokuments, *forme*, *signe* und *medium*,¹⁰ dessen Besonderheiten herausarbeitet und in die Sammlungsprozesse einbezieht. Ohne auf diese Diskussion des digitalen Dokumentes oder Textes näher eingehen zu können,¹¹ kann man im Allgemeinen doch sagen, dass die Umwandlung in eine digitale Form das Buch bzw. allgemeiner den Sammlungsgegenstand¹² entmaterialisiert, ihn hinsichtlich seiner Bedeutungsfunktion dynamisiert und neu kontextualisiert. Alle drei Aspekte stehen für elementare funktionale und epistemologische Verschiebungen, und man kann, obwohl das Wort mittlerweile inflationär ist, zurecht von einem Paradigmenwechsel sprechen. Am deutlichsten tritt er der Bibliothek unter dem Begriff der Entmaterialisierung entgegen, wobei es sich dabei, genau genommen, eher um eine gewandelte Beziehung von Trägermaterial bzw. Substrat und Text handelt, die es mit sich bringt, dass Text und Substrat keine feste Verbindung mehr eingehen und sich die Möglichkeiten des „Kopierens“ extrem erweitert haben. Dabei lassen sich Kopien nicht nur mühelos und weitgehend kostenfrei in beliebiger Zahl herstellen, die „Tradierung“ erfolgt zum

⁶ Vgl. a. Kempf (2014).

⁷ Sommer (1999).

⁸ Oltersdorf und Schmunk (2016) 181. Dort auch weiteres zur Diskussion in den DH.

⁹ Ernst (2018) 379.

¹⁰ Pédauque (2006).

¹¹ S. u.a. Eggert (2005), Caton (2013), Sahle (2013).

¹² Im Folgenden beziehe ich mich maßgeblich auf textliche, zur Publikation bestimmte Dokumente, doch gelten die Überlegungen im Prinzip auch für nicht-textliche Materialien. Bei nicht zur Publikation vorgesehenen Daten gelten Sonderbedingungen.

ersten Mal in der Geschichte der Schriftüberlieferung im Prinzip fehlerfrei. Zugleich geht mit dem Verlust der engen Verbindung mit dem Träger einher, dass die bislang über das Trägermaterial garantierte Authentizität und Stabilität der Texte schwindet und es neuer Mechanismen der Sicherung von Originalität oder besser Authentizität bedarf. Mit der Leichtigkeit der Distribution und ubiquitären Nutzbarkeit der digitalen Kopie im Internet entfallen zugleich die eingespielten Verfahren von verlegerischer Dissemination und bibliothekarischer Archivierung. Weder die klassischen Mechanismen der Subskription noch der Sammlung analoger Medien zur Sicherung der Herstellung und des Zugangs machen länger Sinn, wenn einerseits der Kopier- bzw. Vervielfältigungsprozess praktisch keine Kosten mehr verursacht und andererseits kein Verfügbarkeitsproblem mehr besteht. Die Wissenschaft vermag sich einen alten Traum zu erfüllen: der Zugriff auf Forschungsliteratur ist von überallher zu jeder Zeit möglich. Insofern ist die Forderung nach bedingungslosem Open Access konsequent, denn die Absicht der Wissenschaft ist bequeme Nutzung, Dissemination und Reputation, nicht aber Gewinnerzielung.¹³ So gesehen wird sich Open Access einfach aus der Logik des Mediums im Wissenschaftskontext durchsetzen, selbst wenn die Rückzugsgefechte auf allen Seiten erbittert geführt werden.¹⁴

Mit der freien Bereitstellung im Netz scheint daher der Sammlungsbedarf seitens der Bibliothek verschwunden. Sucht man im Netz nach wissenschaftlichen Beiträgen findet man sie meist problemlos über Google. Der Sinn des bibliothekarischen Sammelns, der darin besteht, an einem bestimmten Ort möglichst alle gewünschten Titel bereitzustellen, scheint obsolet, wenn alle Publikationen in Zukunft Open Access sind und weiterhin Google oder ähnliche Suchmaschinen-Dienstleister die Suche ermöglichen. Die radikalste Konsequenz aus dieser Einsicht hat die Universitätsbibliothek in Utrecht gezogen, die auf einen eigenen Katalog komplett verzichtete, um sich in Zukunft auf die Google-Suche und WorldCat zu verlassen, in der Erwartung, dass über das Netz alle nötigen Ressourcen gefunden werden.¹⁵ Konsequent zuende gedacht, wäre der nächste Schritt, auch die Sammlung dem Netz zu überlassen. Selbst wenn eine gewisse Unruhe beschleichen mag, ob der Zugang zuverlässig und langfristig sichergestellt ist, ob die Suchalgorithmen zumindest von Google der Forschung oder doch eher der Werbung verpflichtet sind, die Haltung, es sein zu lassen, scheint konsequent.

Doch der Begriff der Sammlung schließt schon traditionell mehr ein, als die möglichst umfassende Anhäufung von Dokumenten und Texten und ihre bibliographische Beschreibung bzw. ihr Katalogeintrag, der es erlaubt gezielt einen Titel oder aber gleichartige Titel zu einem Thema zu

¹³ Suber (2013), Eve (2014), Stäcker (2016).

¹⁴ Vgl. den DEAL Prozess: <https://www.projekt-deal.de>.

¹⁵ <https://www.uu.nl/en/news/where-is-the-catalogue>.

finden. Anders als die Ansammlung bildet die bibliothekarische Sammlung - zumindest ihrer Idee nach - ein geordnetes und strukturiertes Ganzes.¹⁶ Kein Mensch würde auf die Idee kommen, das Internet ein strukturiertes Ganzes zu nennen. Es ist, was schon früh bemerkt wurde, keine Bibliothek.¹⁷ Dieser Ordnungsgedanke ist so alt wie die Bibliothek selbst und ist selbst auch immer wieder aus verschiedenen Gründen als konstitutiv für den Begriff der Bibliothek gesehen worden, selbst wenn es postmoderne Einsprüche gegen die geordnete Bibliothek und Befürworter einer ungezügelter Serendipity gegeben hat.¹⁸ Das geordnete Ganze ist selbst Voraussetzung für die Dekonstruktion der Ordnung, mit anderen Worten, ohne ein Mindestmaß an Ordnung gibt es auch keine Serendipity.¹⁹ Die Ordnungsarbeit der Bibliothek beruht, verkürzt formuliert, in zweierlei: Einerseits in der vereinheitlichenden formalen Beschreibung aller Dokumente zum Zwecke der eindeutigen Beschreibung bzw. zum Wiederfinden, andererseits in der inhaltlichen (Sach-)Erschließung zum Zwecke des Findens von Literatur zu bestimmten Themen und Gegenständen. Grundsätzlich ändert sich daran auch im Falle des Medienwandels nichts. Allerdings ändern sich die Anforderungen und Rahmenbedingungen erheblich.

1 Sammlung als Transformation

Auch wenn es noch Gründe für die Produktion gedruckter Bücher gibt - z.B. empfinden viele Nutzer die Lesbarkeit elektronischer Medien aus verschiedenen Gründen als unzureichend²⁰ - ist doch der Medienwandel längst vollzogen. Auch gedruckte Bücher entstehen digital. Es fehlt bislang nur an geeigneten Verfahren, diesen Umstand systematisch zu nutzen, aber auch an Bereitschaft, die digitalen Originale als wesentliche Komponente des Publikationsprozesses zu betrachten. Der Druck ist in diesem Sinne derivativ und im wörtlichen Sinne „Ausdruck“ des digitalen Originals. Digital und analog sind daher kein Antagonismus (wie Knoche und andere nahelegen)²¹, sondern eine sinnvolle Kombination. Ebenso wenig ist es, sozusagen in einer Rückprojektion anmessen, bezüglich der Ära der Druckgeschichte von „Gutenberg terror“ zu sprechen, den das Internet glücklich beseitigt habe.²² Natürlich muss das digitale Dokument, um für den Ausdruck oder andere Präsentationsformen geeignet zu sein, eine bestimmte Form haben, die eine flexible Nachnutzung für die verschiedenen Zwecke, seien sie analog oder digital, ermöglicht. Dazu haben sich bereits in verschiedenen Feldern Standards etabliert, die nach dem auch im Internet vorherrschenden

¹⁶ Sommer (1999).

¹⁷ Zimmer (2000).

¹⁸ Wegmann (2000).

¹⁹ Stäcker (2009).

²⁰ Tuschling (2014) 39f.

²¹ Knoche (2017) 80.

²² Anderson (2011) 211.

Dokumentenmodell DOM, W3C) in XML formuliert wurden, allen voran TEI²³ und JATS²⁴, wobei es auch Versuche gibt, Texte anders, z.B. mit Graphen zu strukturieren.²⁵ Im Bereich der Forschungsdaten begegnet uns größere Vielfalt, wie CSV, JSON, RDF Serialisierungen usw., deren effiziente Nutzbarkeit indes unter zwei Bedingungen steht, einerseits dass sie möglichst standardisiert sind und andererseits dass sie leicht von Algorithmen gelesen und ausgewertet werden können. Audio- und Videoformate sind in dieser Hinsicht anspruchsvoller als Textformate, streng strukturierte Datenformate sind leichter zu verarbeiten als wenig strukturierte dokumentenzentrierte Daten. Ein Sonderfall bilden die vernetzenden Daten, früher oft mit dem Begriff des Hypertext²⁶ verbunden, heute sprachlich und konzeptionell eher im Bereich des *semantic web* angesiedelt. So lassen sich Beziehungen unter verschiedenen Sammlungen, aber auch in Sammlungen vorhandene Entitäten, die mittels Graphen Netzwerke ausbilden, gut mit RDF beschreiben und als Linked Open Data zugänglich machen, wobei, wie gesagt, die Nutzung von RDF nicht darauf beschränkt ist, Sammlungen zu verbinden. Mit RDF können auch die Daten selbst, die in den Sammlungen enthalten sind, wie auch die Metadaten, die die Daten beschreiben,²⁷ abgebildet werden. Eine besondere Herausforderung ist dabei die Entwicklung bzw. Nutzung geeigneter Ontologien.

Im Textbereich sind die auf XML aufbauenden Standards durch entsprechende Schemadateien fixiert. Solche Standards stellen sicher, dass wesentliche Struktur- und Inhaltselemente von Texten durch so genanntes deskriptives Markup²⁸ erfasst werden. Bibliotheken haben sich an der Entwicklung dieser Standards (z.B. TEI for Libraries)²⁹ beteiligt, aber bislang wenig für deren Umsetzung im bibliothekarischen Alltag getan. Obwohl absehbar ist, dass nur auf der Grundlage strukturell aufbereiteter maschinenlesbarer Dokumente das volle Potential des digitalen Mediums ausgeschöpft werden kann, orientieren sich die meisten Bibliotheken immer noch an dem Modell eines Textes oder Mediums, das durch Metadaten (Katalog) erschlossen wird, sehen aber nicht, dass das Medium selbst über die Logik des Archivierens hinaus, eine eigene neue Erschließungsaufgabe enthält. Eine systematische Umstellung würde auch durch den Umstand begünstigt, dass wissenschaftliche Beiträge, insb. in den MINT Fächern, weitgehend einheitlichen Strukturen folgen,³⁰ die leicht mit deskriptivem Markup erschlossen werden könnten, Markup, das es erlauben würde,

²³ <http://www.tei-c.org>.

²⁴ <https://jats.nlm.nih.gov>.

²⁵ S. vor allem Dekker und Birnbaum (2017), zum Themenkomplex „Text as a Graph“.

²⁶ Landow (2006).

²⁷ S. <https://www.w3.org/TR/vocab-dcat>, vgl. a. <https://www.govdata.de>.

²⁸ DeRose et al. (1990).

²⁹ <http://www.tei-c.org/SIG/Libraries/teiinlibraries>.

³⁰ Typisch ist z.B. der Aufbau von Artikeln nach dem IMRAD Prinzip: Introduction, Methods, Results, and Discussion.

bibliographische Metadaten, Überschriften, Abstrakts, Schlagwörter, Entitäten, Zitate, Haupt- und Unterkapitel und vor allem auch Literaturverzeichnisse selektieren, visualisieren und analysieren zu können. Die Zurückhaltung, sich intensiver der Bearbeitung und Transformation der Dokumente selbst zuzuwenden, liegt zum großen Teil daran, dass man sich über Jahre hinweg der Illusion hingegeben hat, dass man mit „digitalen Inkunabeln“ wie PDF dem elektronischen Medium gerecht würde. Doch PDF simuliert nur den Druck und verhindert eine qualifizierte Nachnutzung. D.h. nicht, dass PDF wie der Druck ein Derivat für bestimmte Zwecke sein kann, es eignet sich aber als maßgeblich layoutbasiertes Format grundsätzlich nicht als Primär- bzw. Masterformat, wie die Entwicklung von PDF/A suggeriert. Dieser Fehleinschätzungen folgen viele Bibliotheken und auch Verlage bis heute. Aus der Einsicht, dass das digitale Potential von Dokumenten und Texten nur dann voll genutzt werden kann, wenn es in maschinenlesbarer, mit strukturellem Markup ausgestatteten Form vorliegt, leitet sich eine wichtige „Sammlungsaufgabe“ der Bibliothek ab. Dokumente müssen in digitaler Form mit standardisiertem Markup gesammelt und, sofern der Volltext und deskriptives Markup nicht vorhanden ist, in Volltext umgewandelt und mit deskriptivem Markup ergänzt bzw. aufbereiten werden. Sammeln bedeutet in diesem Sinn Transformieren und Homogenisieren nach etablierten Dokumentenstandards. Wenn Bibliotheken früher Bücher mit Katalogisaten ausstatteten, um sie zugänglich zu machen, müssen Sie heute Dokumentinhalte durch Volltexte und strukturelle Metadaten zugänglich machen. In diesen Strukturen sind auch die bibliographischen Metadaten enthalten. Sie müssen, wenn dort vorhanden, nicht mehr außerhalb der Dokumente erfasst werden. Der Katalog in diesem Sinne wird zum Index, das Buch zum integralen, sich selbst erklärenden Informationsobjekt.

2 Sammlung und Open Access

Das neue Publikationsparadigma der globalen ungehinderten Verfügbarkeit im Open Access unter freien CC Lizenzen (durchgesetzt als wiss. günstig haben sich CC0, CC BY oder CC BY-SA) erlaubt Bibliotheken, ungehindert Kopien anzufertigen.³¹ *Mutatis mutandis* kehren sie nach einem langen Weg, der bis zur Lizenzierung digitaler Werke und „Enteignung“ der Bibliotheken (nicht Eigentumsfähigkeit der Bibliothek an digitalen Medien, die juristisch nicht mehr als Sache betrachtet werden) zu der Situation der gleichsam primordialen Sammlung, der Bibliothek von Alexandria, zurück und sind wieder in der Lage, nach eigenem Ermessen, jetzt in digitalen Skriptorien, zu „kopieren“. Auch von dieser Möglichkeit wird bisher wenig Gebrauch gemacht, vor allem mit

³¹ Natürlich wird es noch eine längere Phase des Übergangs im wissenschaftlichen Publikationswesen geben und die klassischen Kauf- und Lizenzmodelle werden, mit abnehmender Bedeutung noch längere Zeit Bestand haben. Eine Zukunft haben sie aber m.E. nicht.

Verweis auf die Online-Verfügbarkeit der Open Access Medien und mit Verweis auf zentrale Sammelstellen. Doch solche Stellen gibt es für den Nutzer allenfalls virtuell. Weder erbringen die zentralen bibliothekarischen Sammelstellen, allen voran die Nationalbibliotheken, die integrale Bereitstellung digitaler Medien, noch auch leisten sie die notwendige Dokumenttransformation. Das wäre nach der gegenwärtigen Situation vermutlich auch eine Überforderung. In Deutschland wird bezeichnenderweise die einzige Sammlung, die ansatzweise diese Kriterien erfüllt, nicht durch eine Bibliothek, sondern durch ein Forschungsprojekt bereitgestellt, dem Deutschen Textarchiv³², begrenzt allerdings auf historische Quellen, das als Referenzkorpus vor allem für die deutsche Sprachwissenschaft aufgebaut wurde, indes auch darüber hinaus breites Interesse fand. Forscherinnen und Forscher aus dem Bereich der DH, die mit digitale Methoden arbeiten, gehen vor allem dorthin oder anderen Korpora wie Gutenberg³³ oder TextGrid (Zeno)³⁴, nicht aber zur Bibliothek. Angesichts der Situation hat sich in der Forschung schon der Begriff der „opportunistischen Sammlung“³⁵ eingebürgert: nicht systematisch aufbereitete, sondern zufällig verfügbare Sammlungen sind Grundlage der Forschung. Angesichts dessen sollte verstärkt auf das aufbereitende und erschließende Sammeln von Open Access Werken Wert gelegt und Open Access als Chance zu sammeln und nicht als bequeme Erübrigung interpretiert werden.

3 Die Sammlung just in time und just in case oder das Ende der traditionellen Sammlung

Ein zentraler Streitpunkt bei der Umwandlung des DFG Sondersammelsystems in das Fachinformationssystem in Deutschland war die Frage des Wechsel vom “vorausschauenden Bestandsaufbau” (just in case) zur Befriedigung des aktuellen Forschungsbedarfs (just in time).³⁶ Mit der digitalen Wende verlieren allerdings diese Kategorien an Bedeutung. Freie digitale Sammlungen im Open Access sind im Prinzip beliebig aggregierbar. Anders als analoge Medien benötigen digitale Dokumente keinen teuren Magazinraum und nehmen zumindest in Textform kaum Platz ein. Es ist durchaus vorstellbar, dass in Zukunft alle Dokumente einer großen bibliothekarischen Sammlung auf einem Smartphone Platz finden. Da sich die Kosten vom Erwerben (Kauf, Lizenzierung) auf die Herstellung (APC, BPC) verlagern, kann die Bibliothek grundsätzlich eine Sammlung aufbauen, die *just in case* und *just in time* alle Fragen zu beantworten geeignet ist, weil sie kostenfrei aggregierbar ist und “im Prinzip” alle wissenschaftlichen Texte enthalten kann. Der limitierende Faktor liegt allein

³² <http://www.deutschestextarchiv.de>.

³³ <http://gutenberg.spiegel.de>.

³⁴ <https://textgrid.de/digitale-bibliothek>.

³⁵ Schöch (2017).

³⁶ Griebel (2014), Knoche (2015), Kempf (2014).

in der Fähigkeit zur medialen Transformation, integrierenden Homogenisierung und intelligenten Retrievaltechniken. Es liegt auf der Hand, dass angesichts dessen die Bedeutung der Sammlung im Sinne einer Menge von verfügbaren Medien schwindet. Wenn zukünftig alle Texte frei zirkulieren, wenn sich alle Bibliotheken in wahrhaft alexandrinischem Umfang gegenseitig kopieren, dann liegt die Qualität der Sammlung nicht mehr im *Content*, sondern in der Fähigkeit, den Content zu prozessieren, d.h. zu ergänzen, aufzubereiten und sinnvoll zu selektieren, vulgo zu ordnen. Die Sammlung als *Content* ist sozusagen überall. Die lokale Sammlung, die bislang die Bedeutung einer Bibliothek ausmachte, ist so gesehen nicht mehr ihre eigene, sondern ein vergemeinschaftetes zirkulierendes, sich stetig ergänzendes und verbesserendes Gut, das gezielt für die Informationsversorgung der jeweiligen eigenen Klientel genutzt, angepasst oder ergänzt wird. Natürlich ruft eine solche Konstruktion nach Kooperation. Denn die Umwandlung und Aufbereitung der Texte ist ein aufwändiges Geschäft, zumal erst wenige Textproduzenten oder Verlage systematisch digitale Texte mit deskriptivem Markup bzw. in strukturierter Form in geeigneten Standards anbieten.³⁷ So ist einerseits ein systematischer und abgestimmter Prozess der *Datafikation*³⁸ des analogen Schrifttums erforderlich, andererseits die Aufbereitung und Transformation von analogen und digitalen Materialien zu für die digitale Forschung nachnutzbaren Daten- und Dokumentsammlungen. Mit der Transformation des analogen Schrifttums in digitale Form verändert sich auch die Bedeutung der vorhandenen analogen Sammlungen. Das Digitalisat, als Image, Volltext mit und ohne deskriptivem Markup bis hin zur digitalen Edition fungiert als Proxy des Originals und ist als ein Prozess der Erschließung anzusehen. Während früher der Katalog als Proxy zum Original führte und seine Nutzbarkeit sicherstellte, vertritt heute der digitale Proxy das Original und erlaubt erst in dieser Umwandlung moderne Forschung, die sich digitaler Methoden bedient, wie automatisierte Bilderkennung, Text- and Data Mining, Visualisierung, multimodales Edieren, semantische Verknüpfung von Entitäten und Netzwerkanalyse, Ontologieentwicklung, Topic Modelling, Stilometrie, Sentiment Analysis usw. Die Arbeit mit den Dokumenten verlagert sich ins Digitale. D.h. nicht, dass das Original als solches seinen Wert und kulturelle Bedeutung verlöre und in seiner Materialität selbst als Forschungsgegenstand sein Recht forderte, doch erweitern sich mit den neuen Methoden auch die Anforderungen an seine Erschließung im Sinne der *Datafikation*. Der digitale Proxy gehört wie der Katalogeintrag selbst nicht mehr der Sammlung an, sondern unterliegt wie *born digitals*, die sozusagen beide Elemente integrieren, anderen Gesetzmäßigkeiten. Es schöpft das Potential des Digitalen nicht aus oder muss sogar als Einschränkung von

³⁷ S. z.B. SpringerOpen: <https://www.springeropen.com/get-published/indexing-archiving-and-access-to-data/api>.

³⁸ Dabei geht es weniger um Zugänglichkeit, sondern im Sinne der FAIR Prinzipien um die vollständige und uneingeschränkte Nachnutzbarkeit der Objekte in allen vorhandenen Formaten. Ceynova (2015) 273 spricht von einem Wandel vom "Bereitstellen des Zugangs zum Bereitstellen hochwertiger, originaler digitaler Masterdateien".

Forschungsmöglichkeiten betrachtet werden, wenn Bibliotheken in Anhänglichkeit an traditionelle Sammlungskonzepte Digitalisate mit Nutzungsbedingungen, die keine uneingeschränkte Nutzung erlauben, wenn sie nicht den massenhaften Download für *big data* Analysen gestatten, verminderte Bildqualitäten anbieten und so die Datenqualität verschlechtern oder auch Volltexte oder mit deskriptivem Markup versehen Volltext verbergen oder Metadaten nicht über frei verfügbare Schnittstellen zur Verfügung stellen. Nur wenn diese, mit den FAIR Prinzipien zu verbindenden Möglichkeiten gegeben sind und sie sich von dem alten Vorstellungen frei gemacht haben, die die Bedeutung einer Sammlung aus dem Gedanken speist, dass man Sammlungsgegenstände besitzt, die andere nicht haben, erfüllen Bibliotheken ihren Informationsauftrag im digitalen Raum. Man muss die Strategie umkehren: Nur wenn eine Bibliothek *kein* Dokument besitzt, das nur bei ihr zu finden ist, hat sie ihre Bestimmung vollständig erfüllt! Diese Strategie ist als dauerhafter Prozess zu verstehen: unausgesetzt alles von überallher zu versammeln und zugleich das, was selbst erzeugt und bereitgestellt wird, überallhin zu verteilen. Je besser dies gelingt, umso besser die Bibliothek. Ihr neues Alleinstellungsmerkmal ist die Effizienz in den genannten Transformationsprozessen und in der Verbesserung der grundsätzlich nicht ortsgebundenen digitalen Sammlung bzw. der imaginären einen großen Sammlung, die alle teilen, sei sie eine redundante Kopie, sei sie ein distributives semantisches Netz. Mit Blick auf diese Entwicklungen haben Bibliothekstheoretiker wie Anderson zwar das Richtige gesehen: “and it seems highly likely that the very idea of the ‘collection’ will be overhauled if not obviated over the next ten years, in favor of more dynamic access to a virtually unlimited flow of information”.³⁹ Zugleich springen sie aber zu kurz, wenn sie die Rolle und Optionen der Bibliothek in der Ermöglichung dieses “flow of information” unterschätzen, darin dass sie einseitig die negativen, nicht aber die positiven Seiten des digitalen Wandels und die sich bietenden neuen Chancen bei der Sammlungsbildung betonen.

Als Fazit bleibt festzuhalten: Die Sammlung lebt und Sammeln bleibt ein konstitutives Element bibliothekarischer Arbeit. Sammeln bedeutet Erwerben, Erschließen und Benutzen. An diesen Grundfunktionen ändert sich auch in digitalen Zeiten nichts. Was sich ändert, ist - und das ist nichts Geringes -, dass die Sammlung kein statischer Bestand mehr ist. Sammeln dient nicht mehr vorrangig dem Behalten und Bewahren, sondern vor allem dem erschließenden Aufbereiten, Ordnen und Teilen. Die digitale Sammlung gehört nicht mehr der einen Bibliothek und ist nicht mehr ihr distinktives Merkmal, sondern sie gehört gleichermaßen allen Bibliotheken. Die Kunst des Sammelns besteht in der Daten- und Dokumentaufbereitung, in der Schaffung von forschungsintegralen

³⁹ Anderson (2011) 215.

Datenzyklen, kurz, in der *Datafication* und der Entwicklung der Bibliothek als Schnittstelle.⁴⁰ Der unerschöpfliche Raum des Internet liefert den Rohstoff, der in Bibliotheken zusammengeführt, für den wissenschaftlichen digitalen Gebrauch veredelt und wieder zurückgespielt wird. Die Sammlung ist tot, es lebe die Sammlung!

Literaturverzeichnis

- Anderson, Rick (2011): Collections 2021: the future of the library collection is not a collection. In: *Serials* 24, 3, 211-15.
- Caton, Paul (2013): On the term 'text' in digital Humanities. In: *Literary and Linguistic Computing*, 28, 209–20.
- Ceynova, Klaus (2015): Vom Wert des Sammels und vom Mehrwert des Digitalen - Verstreute Bemerkungen zur gegenwärtigen Lage der Bibliothek. In: *BIBLIOTHEK – Forschung und Praxis*, 39 (3), 268–76.
- Dekker, Ronald; Birnbaum, David J. (2017): It's more than just overlap: Text As Graph. Presented at Balisage: The Markup Conference 2017, Washington, DC, August 1 - 4, 2017. In: *Proceedings of Balisage: The Markup Conference 2017*. Balisage Series on Markup Technologies, 19 . <https://doi.org/10.4242/BalisageVol19.Dekker01>.
- DeRose, Steven J.; Durand, David G.; Mylonas, Elli; Renear, Allen H. (1990): What Is Text, Really? In: *Journal of Computing in Higher Education*, 1, (Dezember) 2, 3–26. <https://doi.org/10.1007/BF02941632>.
- Eggert, Paul (2005): Text-encoding, Theories of the Text, and the 'Work-Site'. In: *Literary and Linguistic Computing*, 20 (4), 425–35.
- Ernst, Wolfgang (2018): Die Unwahrscheinlichkeit von Wissenstradition und die Beharrlichkeit der Bibliothek. In: *BIBLIOTHEK – Forschung und Praxis*, 42, (2), 379–86.
- Eve, Paul Martin (2014): Open Access and the Humanities, Cambridge. Verfügbar unter <http://www.cambridge.org/doi/titles/open-access-and-humanities-contexts-controversies-and-future>.
- Griebel, Rolf (2014): Ein 'folgenreicher' Paradigmenwechsel. Die Ablösung der Sondersammelgebiete durch die Fachinformationsdienste für die Wissenschaft. In: *ZfBB*, 61, (3), 138–57.
- Kempf, Klaus (2014): Bibliotheken ohne Bestand? In: *BIBLIOTHEK – Forschung und Praxis*, 38, (3), 365–97. <https://doi.org/10.1515/bfp-2014-0057>.
- Knoche, Michael (2015): Haben wissenschaftliche Bibliotheken noch einen Sammelauftrag? In: *o-bib*, 4, 78–83.
- Knoche, Michael (2017): Die Idee der Bibliothek und ihre Zukunft. Göttingen.
- Landow, George P. (2006): Hypertext 3.0 : critical theory and new media in an era of globalization.
- Lankes, R. David (2016): The Atlas of New Librarianship. Cambridge [Mass.] et al.
- Lankes, R. David (2018): Why do we need a new library science. In: *BIBLIOTHEK – Forschung und Praxis*, 42 (2), 338–43.
- Oltersdorf, Jenny; Schmunk, Stefan (2016): Von Forschungsdaten und wissenschaftlichen Sammlungen. In: *BIBLIOTHEK – Forschung und Praxis*, 40 (2). Verfügbar unter <https://doi.org/10.1515/bfp-2016-0036>.

⁴⁰ Stäcker (i. Druck).

- Pédauque, Roger T. (2006): *Le document à la lumière du numérique*. Herausgegeben von Jean-Michel Salaün. Caen: C & F Éd.
- Sahle, Patrick (2013): *Digitale Editionsformen : zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels*. 3 Bde. Norderstedt.
- Schöch, Christof (2017): Aufbau von Datensammlungen. In: *Digital Humanities*, hg. v. Jannidis F., Kohle H., Rehbein M. Stuttgart, 223–33.
- Sommer, Manfred (1999): *Sammeln: ein philosophischer Versuch*. Frankfurt am Main.
- Stäcker, Thomas (2009): Rez.: Nikolaus Wegmann, *Bückerlabyrinth*. Suchen und Finden im alexandrinischen Zeitalter. In: *Arbitrium* 12 2, 265–68. Verfügbar unter <https://doi.org/doi:10.1515/arbi.2003.21.3.265>.
- Stäcker, Thomas (2016): Neue Entwicklungen im wissenschaftlichen Publikationswesen : oder warum brauchen geisteswissenschaftliche Bibliotheken Open Access. In: *„Ei, dem alten Herrn zoll“ ich Achtung gern“* : Festschrift für Joachim Veit zum 60. Geburtstag, 703–21.
- Stäcker, Thomas (i. Druck): *Literaturwissenschaft und Bibliothek – eine Beziehung im digitalen Wandel*. Metzler.
- Suber, Peter (2013): *Open Access*, Cambridge Mass. Verfügbar unter <https://mitpress.mit.edu/books/open-access>.
- Tuschling, Jeanine (2014): *Bestandsaufbau für digitale Publikationen an Forschungsbibliotheken*. Ein exemplarisches Erwerbungskonzept. Berlin. (Berliner Handreichungen zur Bibliotheks- und Informationswissenschaft, 383).
- Wegmann, Nikolaus (2000): *Bücherlabyrinth : Suchen und Finden im alexandrinischen Zeitalter*. Köln [u.a.].
- Zimmer, Dieter E. (2000): *Die Bibliothek der Zukunft: Text und Schrift in den Zeiten des Internet*. Hamburg.



Prof. Dr. Thomas Stäcker
Direktor der Universitäts- und Landesbibliothek Darmstadt
Magdalenenstr. 8
D-64289 Darmstadt
direktion@ulb.tu-darmstadt.de